Sensimetrics Series in Human Communication

Speech Production and Perception I

Workbook

© 1994, 1995, 1996, 1997 Sensimetrics Corporation

Contents

Spectrograms

Why We Make Spectrograms 4

Going from Spectra to a Spectrogram 5

Finding Formants 6

Bandwidths 7

How We Use Spectrograms 8

Vowel Acoustics

Basic Properties 9

Measuring Formant Frequencies 10

Articulatory and Acoustic Descriptions 13

Constructing F_1/F_2 Plots 15

Physical Differences between Speakers 18

Variation in Articulation 21

Consonant Acoustics

Aerodynamic Effects 25

Degree of Constriction 26

Place of Articulation 27

Voicing 28

Nasalization 29

Speech Perception

What is an Acoustic Cue? 30

An Identification Task 31

Identification Experiments 32

Discrimination Experiments 34

Modes of Perceptions 36

Explaining Categorical Perception 37

Testing the Theories 39

Child and Adult Acquisition 43

Vowel Perception

Phonetic and Perceptual Information 47

Normalization 49

Auditory Processing 51

Auditory Experiments 52

Acknowledgments *53*

Spectrograms: Why We Make Spectrograms

1. What characteristic of speech makes it difficult to analyze using single spectra?

2. Why is it difficult to analyze speech by examining its changing (short-term) spectrum?

3. What are the three dimensions of a display that represent the sequence of speech sounds?

Spectrograms: Going from Spectra to a Spectrogram

1. What do the different shades of gray in a spectrogram represent? What do the different shades correspond to in a spectrum?

2. How are the frequencies of the components of a sound represented in a spectrum? What part of a spectrogram represents these frequencies?

3. How is the sequence of sounds in an utterance represented in a spectrogram? Where are the sounds at the beginning of an utterance represented? The end?

4. How are differences in amplitude represented in a sound spectrum? In a spectrogram?

5. How are different frequencies represented in a sound spectrum? In a spectrogram?

6. How are different times represented in a sound spectrum? In a spectrogram?

Spectrograms: Finding Formants

1. How many formants can you distinguish in a spectrogram of a typical vowel? Can you easily distinguish this many formants in a back vowel like /u/ or /ɔ/?

2. What property of diphthongs is clearly visible in spectrograms? Why do the spectrograms of diphthongs look different from spectrograms of simple vowels (monophthongs)?

3. In what range of frequencies do you expect to find the first formant (F_1) in adult vowels?

4. In what range of frequencies do you expect to find the second formant (F_2) in adult vowels?

5. In what range of frequencies do you expect to find the third formant (F_3) in adult vowels?

6. In what range of frequencies do you expect to find the fourth formant (F_4) in adult vowels?

7. What differences do you see between the spectrograms of the adult male sample vowels and the adult female sample vowels?

Spectrograms: Bandwidths

Spectra can be calculated using different "windows" of a waveform. Changing the duration of the window changes the characteristics of the spectrum.

1. Why does a long window allow more precise estimates of the component frequencies of a sound?

2. Why does a short window allow more precise estimates of the times of events in a sound?

3. Why does a long window make estimates of the times of events in a sound less precise?

4. Why does a short window make estimates of the frequencies of the components of a sound less precise?

5. What features of a sound spectrum are emphasized by using a long window? What does a spectrogram calculated using a long window emphasize? What kind of spectrogram is this?

6. What features of a sound spectrum are emphasized by using a short window? What does a spectrogram calculated using a short window emphasize? What kind of spectrogram is this?

Spectrograms: How We Use Spectrograms

1. Spectrograms can be used to measure various properties of vowels. What properties of vowels can be measured using spectrograms?

2. Spectrograms can also be used to measure various properties of consonants. What properties of consonants can be measured using spectrograms?

3. Spectrograms can be used to measure properties of sounds other than speech. What kinds of sounds can be measured using spectrograms? What kinds of sounds do you think might be profitably investigated using spectrograms and why?

4. What are the differences that spectrograms show between the middle consonants in words like "sober" and "soaper"? What other differences between these words do spectrograms show? Can you think of another pair of words that might have similar differences?

Vowel Acoustics: Basic Properties

Listen to the source sounds. Select a filter and listen to the output sound. Try other filters or try adjusting the filter and listen to the output sound again.

1. Can filters change the loudness of the output sound? Do all filters make the output sound softer? Louder? Which ones?

2. Can the filters shown in this topic change the pitch of the output sound? Do any of the filters make the output sound higher-pitched than the input? Lower-pitched? Which ones?

3. Can the filters shown in this topic change the quality or "perceptual color" of the output sound? Do any of the filters make the output sound "brighter" or "duller" than the input sound? Do they make the output sound "muffled" or "sharp"? Which ones?

4. Do any of the filters produce vowel-like output sounds? Which filters? Which vowel sounds? What does the output of the example of the four-peak filter sound like?

Vowel Acoustics: Measuring Formant Frequencies

Measure the formant frequencies of your pronunciations of the vowels /i, ɪ, ɛ, æ, ɑ, ɔ, ʊ, u/.

1. What are the formant frequencies you measured in your own pronunciations of these vowels?

Vowel	*F1*	*F2*	*F3*
/ i /			
/ ɪ /			
/ ɛ /			
/ æ /			
/ ɑ /			
/ ɔ /			
/ ʊ /			
/ u /			

2. Are your formant frequencies in /i/ the same as the formant frequencies in the pre-recorded sample of /i/? What about in /u/? In /ɑ/? Are your formant frequencies all higher than the formant frequencies in the pre-recorded vowels? Lower?

3. What are the formant frequencies you measured in these vowels in the problem set?

Vowel	*F1*	*F2*	*F3*
/ i /			
/ ɪ /			
/ ɛ /			
/ æ /			
/ ɑ /			
/ ɔ /			
/ ʊ /			
/ u /			

4. Diphthongal vowels are usually transcribed as a sequence of two simple vowels. Are they really acoustically like a sequence of two monophthongs? What else occurs in the middle of the vowel? Are the formants of the diphthongs really like the formants of the monophthongs that supposedly compose them?

5. What are the formant frequencies at the beginnings and ends of your pronunciation of these diphthongs?

	start			*end*		
Vowel	F_1	F_2	F_3	F_1	F_2	F_3
/ eɪ /						
/ ɪʊ /						
/ ɔʊ /						
/ ɔɪ /						
/ ɑɪ /						
/ æʊ /						

6. Are there any differences between your diphthongs' formant frequencies and the formant frequencies of the pre-recorded vowels? Is the pattern of the differences similar to the differences between your monophthongs and the pre-recorded ones?

7. What are the formant frequencies at the beginnings and ends of these example diphthongs?

	start			*end*		
Vowel	F_1	F_2	F_3	F_1	F_2	F_3
/ eɪ /						
/ ɪʊ /						

/ ɔʊ / _____ _____ _____ _____ _____ _____

/ ɔɪ / _____ _____ _____ _____ _____ _____

/ ɑɪ / _____ _____ _____ _____ _____ _____

/ æʊ / _____ _____ _____ _____ _____ _____

Vowel Acoustics: Articulatory and Acoustic Descriptions

Move the blue control handle around in the vowel quadrilateral. Observe the position of the model tongue and listen to the synthetic vowel that corresponds to that tongue position.

1. In an /i/, the tongue is raised close to the hard palate. This enlarges the pharyngeal cavity and narrows the oral cavity. What type of resonant cavity resembles the shape the vocal tract has while making an /i/?

2. Does this kind of resonant cavity usually have a high-frequency resonance or a low-frequency resonance in speech? What formant frequency probably corresponds to this resonance?

3. As you lower tongue and jaw from /i/, what sequence of vowels (or vowel-like sounds) do you hear? How does the first formant frequency change?

4. As you lower the position of the tongue in a vowel, what happens to the size of the pharyngeal cavity relative to the oral cavity? How would you expect this change to affect the resonant frequency of the pharyngeal cavity? Is this the change that you see in the sequence of vowels as the tongue lowers?

5. In an /u/, the inferior part of the pharyngeal cavity is large, and the tongue constricts the top of the pharyngeal cavity, leaving a small opening into the oral cavity. The oral cavity is fairly large, but the rounded lips make a narrow opening. What kind of resonant cavities have these shapes?

6. As you lower the tongue and jaw from an /u/, what happens to the size of the pharyngeal cavity? How would this change their formant frequencies? Is this what you observe in the formant frequencies of the synthetic vowels?

7. In /ɯ/ and /u/, the oral cavity is fairly long, since the tongue is pulled posteriorly. As you move the tongue forward toward /i/ and /y/ respectively, what happens to the length of the oral cavity? What acoustic change would you expect this articulatory change to cause?

8. As you move the model tongue forward to produce an /i/ or /y/, what formant frequencies do you see changing? Does this acoustic change correspond to the change that you expect to see based on the change in the length of the oral cavity?

9. What happens to the oral opening of the vocal tract when the lips are rounded? What acoustic effect do you expect this change to have? How does rounding the lips affect the formant frequencies of a vowel?

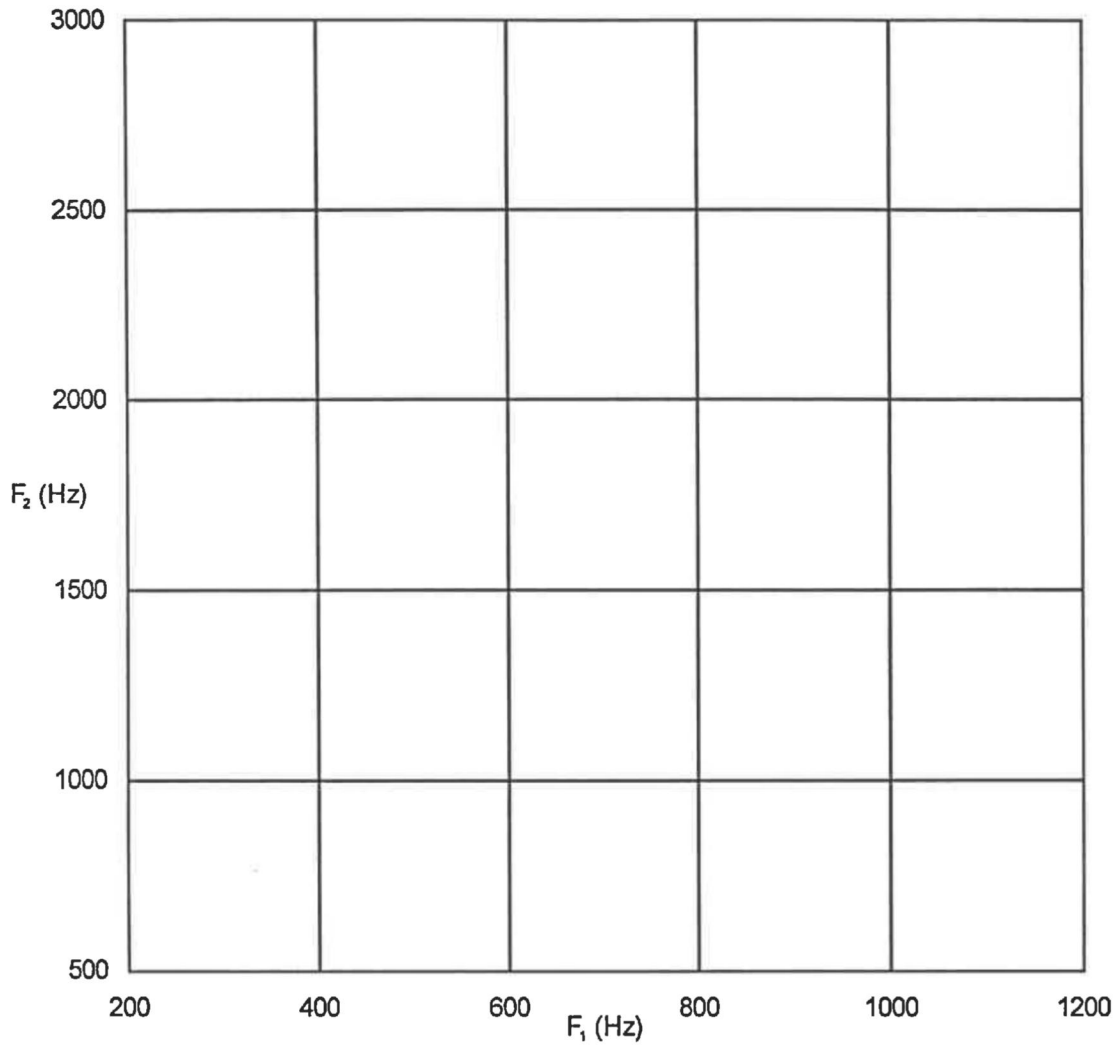
Vowel Acoustics: Constructing F_1/F_2 Plots

Measure and plot the formant frequencies of vowels from the model vocal tract. Compare them to a plot of the formant frequencies of your own vowels. Do the two plots have similar patterns?

1. What are the frequencies measured from the frequency response of the model vocal tract?

Vowel	F_1	F_2	*Vowel*	F_1	F_2
/ i /	_____	_____	/ u /	_____	_____
/ ɪ /	_____	_____	/ ʊ /	_____	_____
/ e /	_____	_____	/ o /	_____	_____
/ ɛ /	_____	_____	/ ɔ /	_____	_____
/ æ /	_____	_____	/ ɑ /	_____	_____

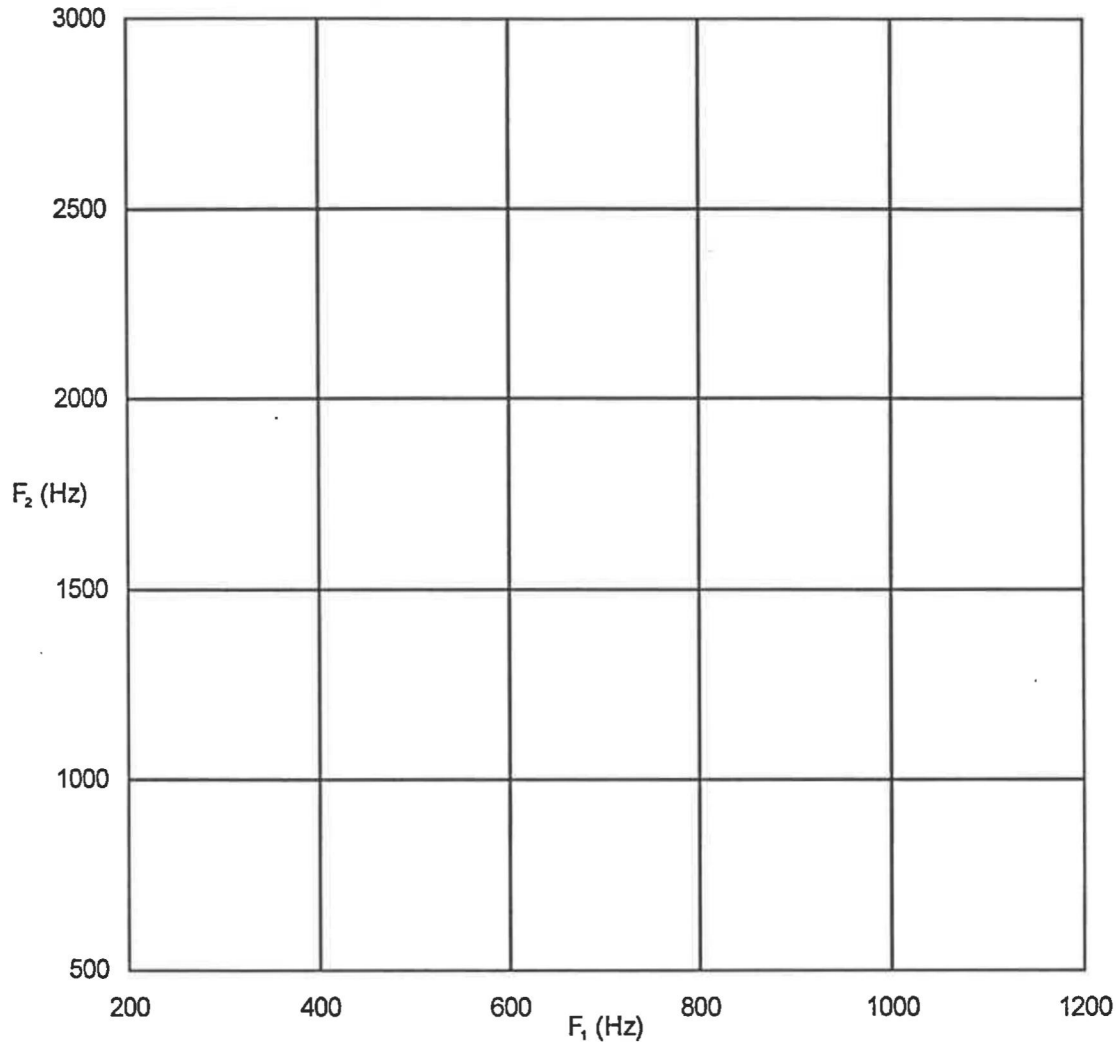
2. Make an F_1/F_2 plot of your measurements by putting the phonetic symbol for each vowel on the graph below.



3. What are the formant frequencies measured from your own productions of these vowels?

Vowel	*F1*	*F2*	*Vowel*	*F1*	*F2*
/ i /			/ u /		
/ ɪ /			/ ʊ /		
/ e /			/ o /		
/ ɛ /			/ ɔ /		
/ æ /			/ ɑ /		

4. Make an F_1/F_2 plot of the measurements from your own vowels by plotting the phonetic symbol for each vowel on the graph below.



5. When you turn the F_1/F_2 plot of vowels so that the vowels are arranged in a pattern similar to the pattern in the phonetic vowel quadrilateral, which formant corresponds to vowel height? What values of this formant's frequency are found in high vowels? Low vowels?

5. When you turn the F_1/F_2 plot of vowels so that the vowels are arranged in a pattern similar to the pattern in the phonetic vowel quadrilateral, which formant corresponds to vowel backness? What values of this formant's frequency are found in front vowels? Back vowels?

Vowel Acoustics: Physical Differences between Speakers

Compare the formant frequencies in your vowels with a friend's. Compare the formant frequencies in the sample adult vowels with the formant frequencies in the sample child vowels.

1. Measure the first two formant frequencies of each of these vowels from two different speakers.

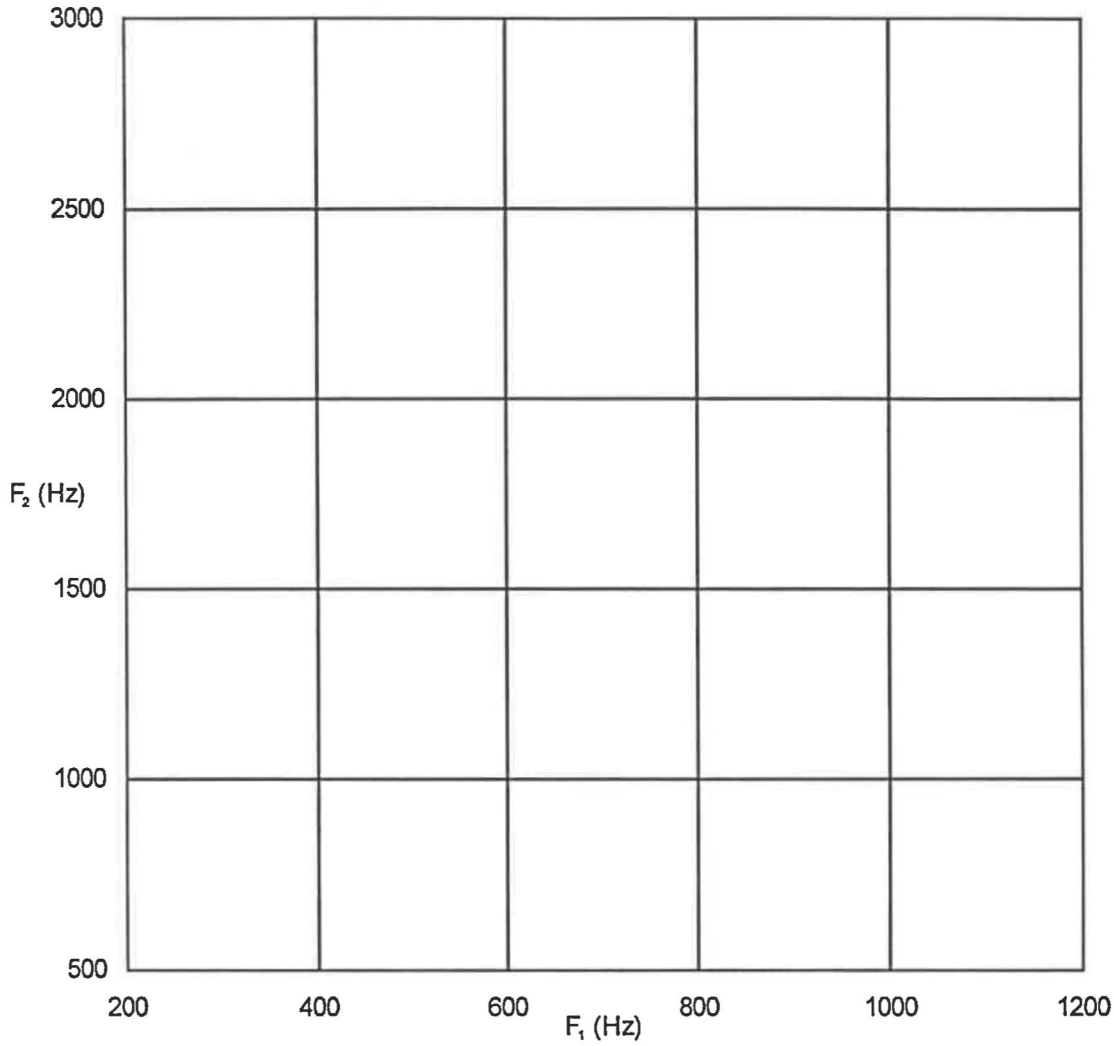
Speaker: _____

Vowel	*F_1*	*F_2*
/ i /		
/ ɪ /		
/ e /		
/ ɛ /		
/ æ /		
/ ɑ /		
/ ɔ /		
/ o /		
/ ʊ /		
/ u /		

Speaker: _____

Vowel	*F_1*	*F_2*
/ i /		
/ ɪ /		
/ e /		
/ ɛ /		
/ æ /		
/ ɑ /		
/ ɔ /		
/ o /		
/ ʊ /		
/u/		

2. Make an F_1/F_2 plot of your measurements by putting the phonetic symbol for each vowel on the graph below. Use a different color for each of the speakers.



3. Compare the two speakers' formant frequencies in /i/, /ɑ/, and /u/. What differences do you see? Is there a consistent proportional difference between the two speakers' formant frequencies? Which speaker has higher formant frequencies?

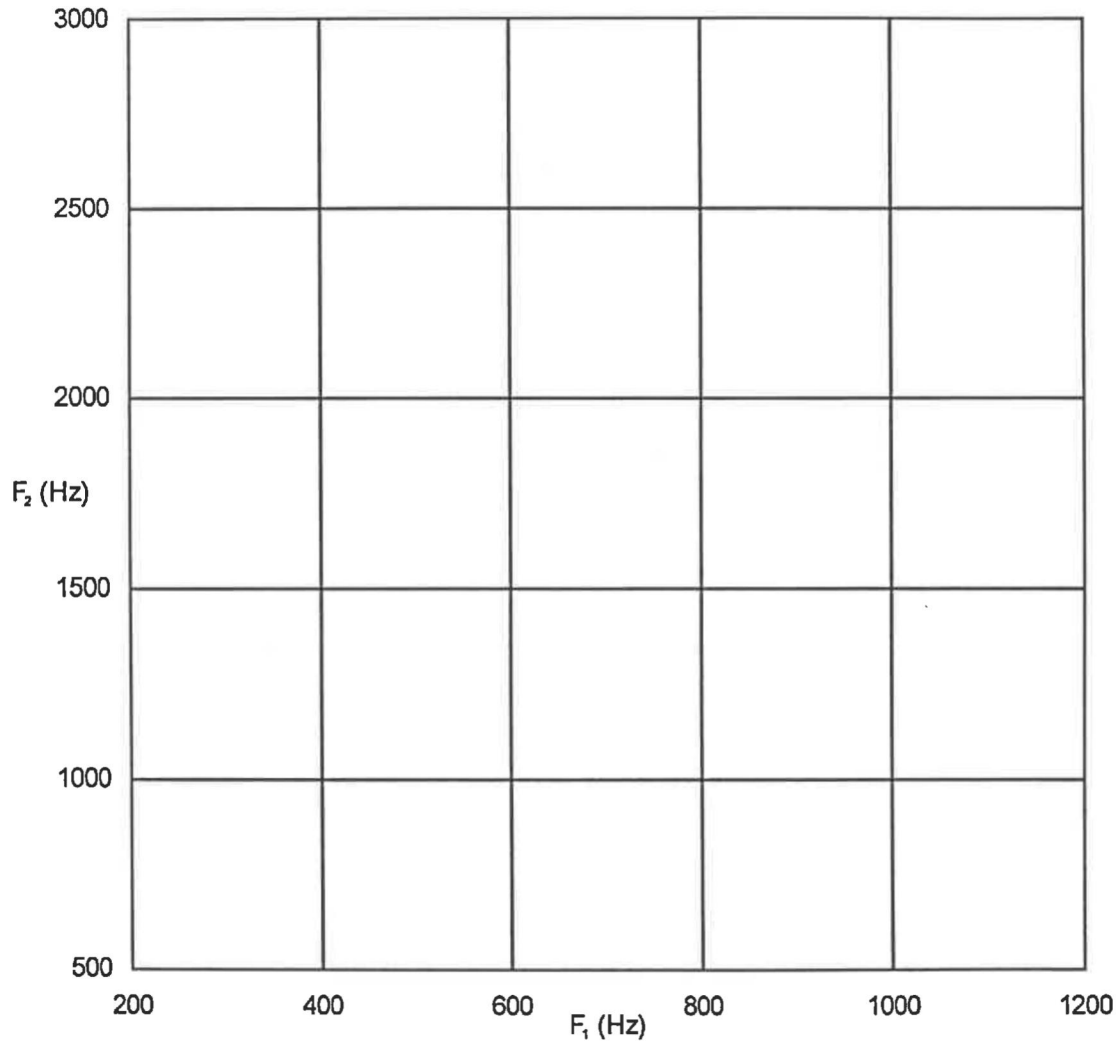
4. What difference in vocal tract size can you infer from the differences between the formant frequencies of your two speakers?

5. Why might it be that two speakers would have a similar F_1 value in /i/ and /u/ despite the fact that they have different vocal tract sizes?

Vowel Acoustics: Variation in Articulation

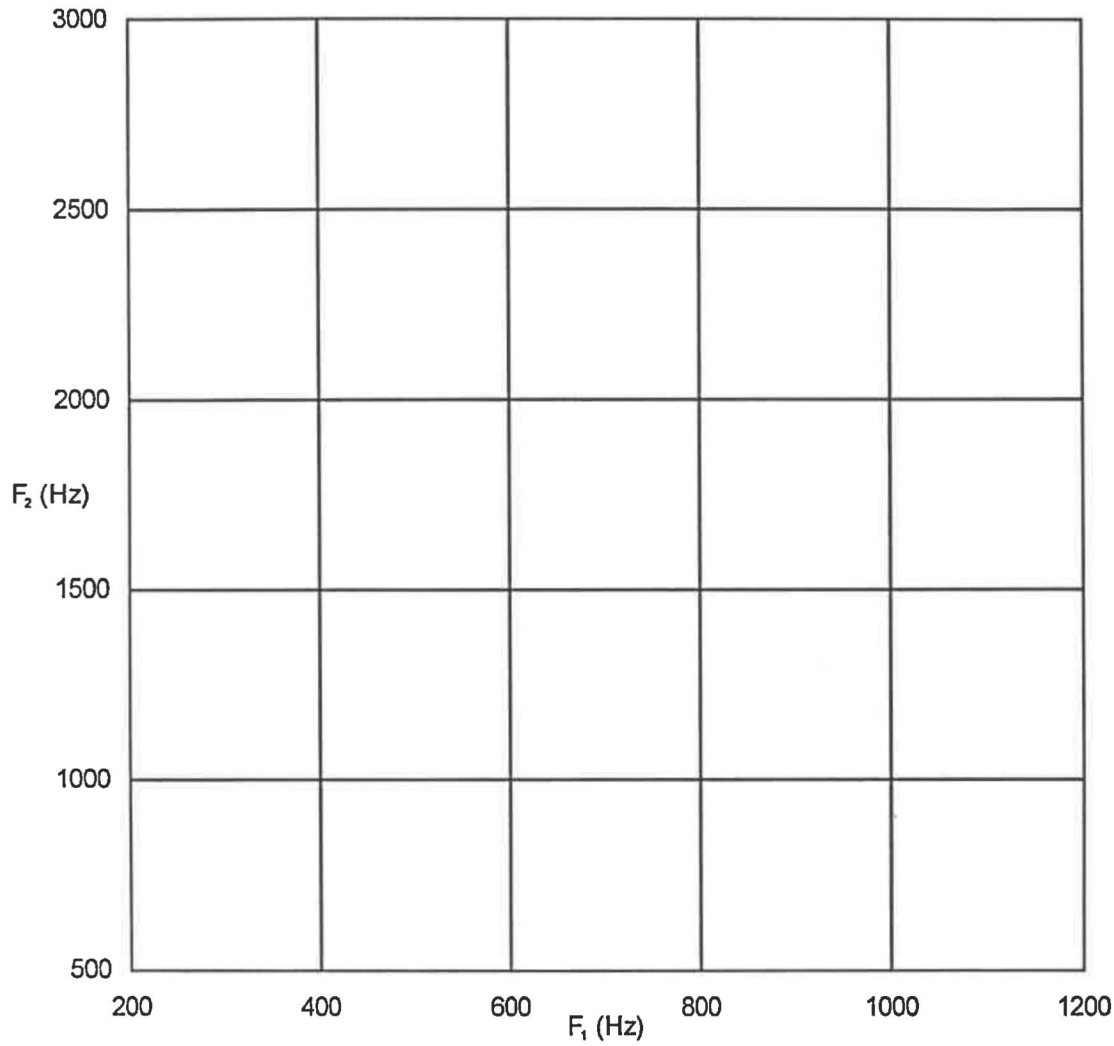
The samples are the words "heed", "had", "hod", and "who'd", pronounced at two different speeds. Compare the slow speech and fast speech variants of the words.

1. Measure the formant frequencies of the sample vowels in each word at the different speech rates and make an F_1/F_2 plot of them. Use different symbols for the vowels at different speech rates (e.g., a triangle for slow, and a circle for fast). Label each group of symbols with the phonetic symbol for the vowel.



2. In this F_1/F_2 plot, are the vowels' formant frequencies changed in any systematic way by the speed with which they were spoken? How?

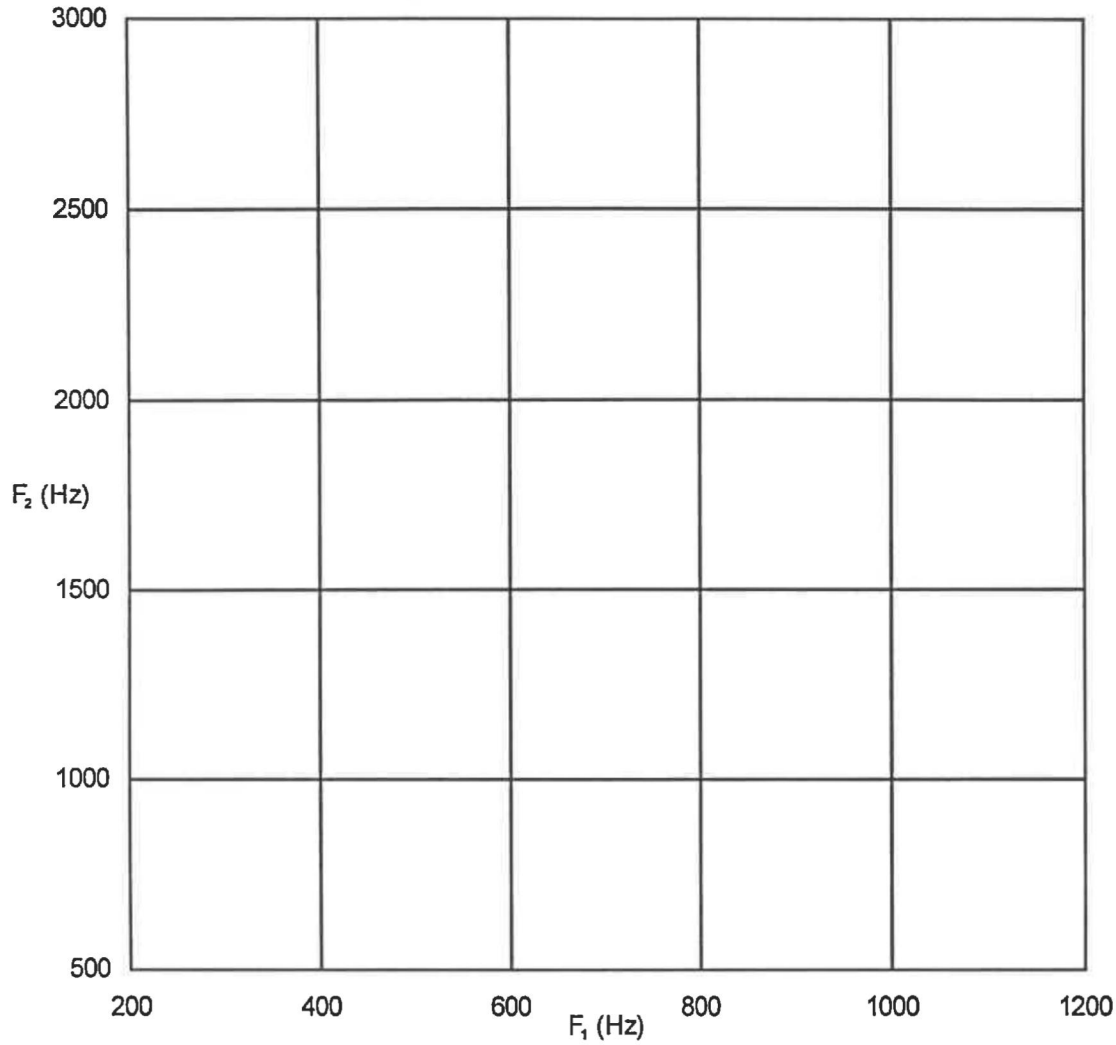
3. Measure the vowel formant frequencies in your pronunciations of the words at different speeds. Make an F_1/F_2 plot of them.



4. In this F_1/F_2 plot, are your vowels' formant frequencies changed systematically by the speed with which the words were spoken? How?

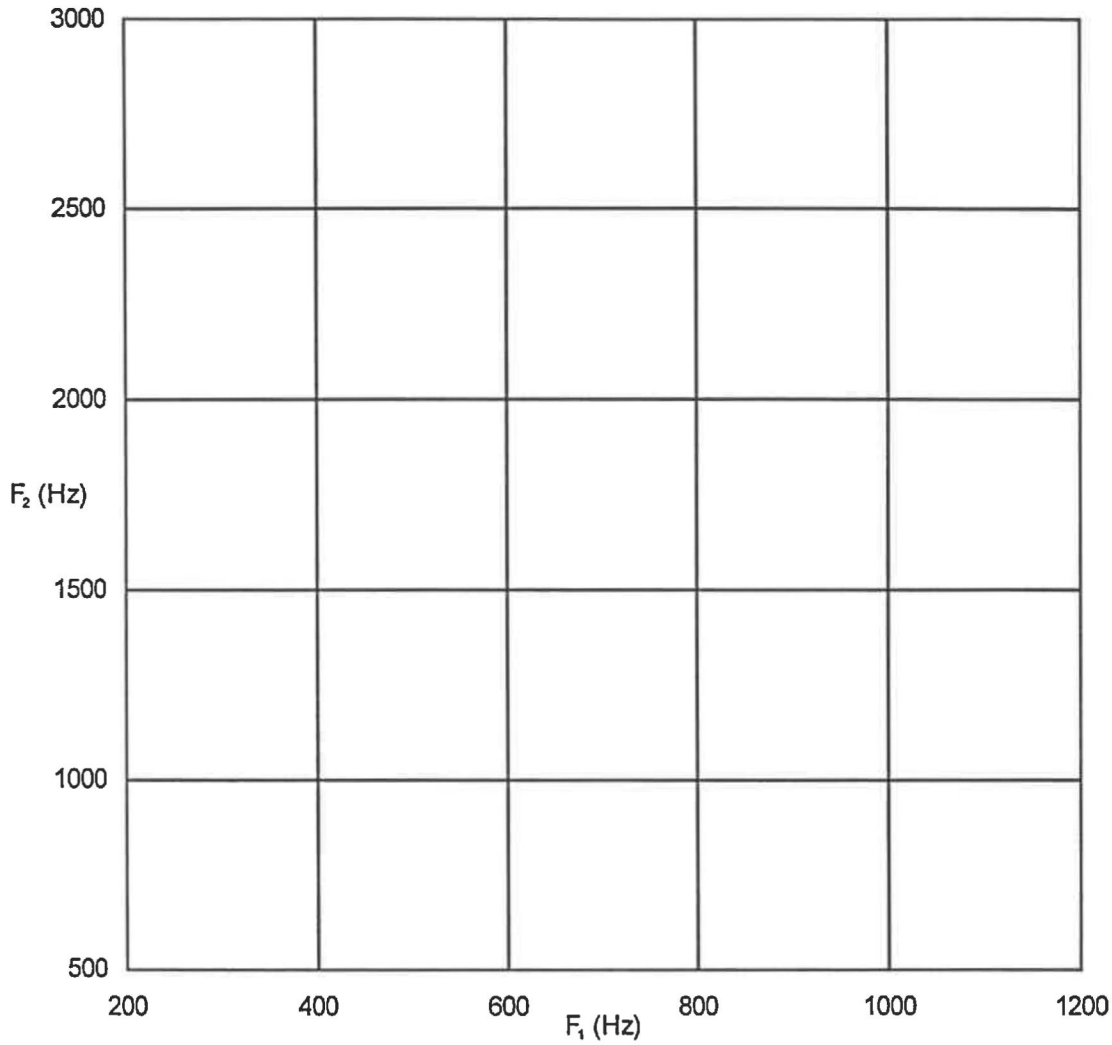
The samples are of the words or word-like utterances "beeb", "bob", "boob", "deed", "dod", "dude", "geeg", "gog", and "goog".

5. Measure the formant frequencies from the center of each of the vowels and make an F_1/F_2 plot of them. Use different symbols for the vowels from words with different consonants (e.g., a circle for /b/, a triangle for /d/, and a square for /g/). Label each group of symbols with the phonetic symbol for the vowel.



6. Do you see any systematic difference between the vowels from the words with different consonants? Do you see a coarticulatory effect in this F_1/F_2 plot? What is it?

7. Measure the formant frequencies from the centers of the vowels in your pronunciations of these words and make an F_1/F_2 plot of them.



8. Do you see any systematic difference between the vowels from the words with different consonants? Do you see a coarticulatory effect? What is it?

Consonant Acoustics: Aerodynamic Effects

1. Why does air flow out of the lungs through the vocal tract?

2. What is the effect of a stop consonant on the airflow out of the vocal tract? What is its effect on the air pressure in the vocal tract?

3. What is the effect of a fricative on the airflow out of the vocal tract? What is its effect on the air pressure in the vocal tract?

3. What is the effect of a nasal consonant on the airflow out of the vocal tract? What is its effect on the air pressure in the vocal tract?

4. How does the air pressure in the vocal tract change during a voiceless stop consonant? During a voiced stop consonant? What differences are there and why?

5. How does the air pressure in the vocal tract change after a stop consonant? How does the airflow through the vocal tract change? Are there differences between voiced and voiceless consonants?

Consonant Acoustics: Degree of Constriction

1. Speech sounds range from those having an unconstricted vocal tract to those having a completely closed vocal tract. What are the classes of phonemes in this range?

2. Consonant articulations change the formant frequencies of the vocal tract, thus creating formant transitions. Which kinds of consonants change the vocal tract's formants the most? The least?

3. What kind of consonant is characterized by a brief period of silence or near-silence? Why?

4. What kind of consonant is characterized by a constriction that produces turbulent airflow?

5. What kind of consonant is characterized only by its formant transitions?

Consonant Acoustics: Place of Articulation

1. Why are formant transitions found both before and after consonants?

2. What formant transition is characteristic of all consonants?

3. What formant transitions are similar in both bilabial and dental/alveolar consonants?

4. What formant transitions are similar in both bilabial and velar consonants?

5. What formant transitions are similar in both dental/alveolar and velar consonants?

6. What differences are there in the transitions for dental/alveolar consonants adjacent to /i/ and adjacent to /ɑ/? What effect does the difference in the vowels' F_2 have on the consonants' formant transitions?

7. What differences are there in the F_1 transitions of consonants adjacent to /i/ and adjacent to /ɑ/? What effect does the difference in the vowels' F_1 have on the consonants' formant transitions?

Consonant Acoustics: Voicing

1. What is the effect of opening the glottis during a consonantal articulation? Which kinds of consonants have an open glottis?

2. What happens when the glottis is still open after a consonantal articulation is released? What kind of consonant is produced?

3. What happens when the glottis closes at the same time as the release of the consonantal articulation? What kind of consonant is produced?

4. What happens if the glottis never opens during the consonantal articulation? What kind of consonant is produced?

Consonant Acoustics: Nasalization

1. What is the effect of lowering the velum during a consonant?

2. What happens when the velum is lowered before the closure of the consonantal articulation?

3. Why is the voicing in a nasal consonant relatively loud, compared to the voicing in a stop consonant?

4. What differences are there between the articulations of /b/ and /m/? /d/ and /n/? /g/ and /ŋ/?

5. Different vowels have different formant frequencies. Does /m/ between different vowels have different formant frequencies? /n/? Why?

Speech Perception: What is an Acoustic Cue?

1. What acoustic differences are there between the examples of the words "pea", "tea", and "key" shown in the spectrograms?

2. Does changing some of the acoustic parts of each word make a difference to how it is perceived? Which changes affect how the words are perceived?

3. What acoustic differences are there between the examples of the words "sue" and "shoe" shown in the spectrograms?

4. Does changing some of the acoustic parts of the words "sue" and "shoe" make a difference to how they are perceived? Which changes affect how the words are perceived?

Speech Perception: An Identification Task

1. Play each of the words in a stimulus continuum. What does the first one sound like? What does the last one sound like?

2. If the words in the stimulus continuum do not all sound the same, which ones sound like the first word in the continuum and which ones sound like the last? Do any of them sound like anything else?

Speech Perception: Identification Experiments

1. Play each of the synthetic words in the first stimulus continuum. What do they sound like? Do they all sound the same?

2. Look at the results from your first experiment. Which stimulus sounded like "ba" most often? Which sounded like "ba" least often?

3. Which stimuli in the continuum sounded like "ba" more than half the time? Which stimuli sounded like "ba" less than half the time? Where does this place the phoneme boundary?

4. Look at the results from your second experiment. Which stimulus sounded like /ɛ/ most often? Which sounded like /ɛ/ least often?

5. Which stimuli in the continuum sounded like /ɛ/ more than half the time? Which stimuli sounded like /ɛ/ less than half the time? Where does this place the phoneme boundary?

6. Look at the identification functions from your two experiments. Do both of the identification functions show that the perceived speech sound varied as the acoustic parameter in the stimulus continuum varied?

7. Does one of the identification functions slope more gradually than the other near the phoneme boundary? Which one? What does that indicate about the way in which the perceived speech sound varied as the acoustic parameter varied?

8. Are the shapes of the "typical" identification functions similar? Which one is steeper near the phoneme boundary? How would you describe the difference between the two identification functions?

Speech Perception: Discrimination Experiments

1. Look at the results of your first experiment. Are all pairs of stimuli (pairs of "ba"s and "da"s) equally easy (or difficult) to tell apart?

2. Which pair or pairs of stimuli were easiest to tell apart? Which pair or pairs of stimuli were most difficult to tell apart?

3. What do the pairs of stimuli that are easy to tell apart sound like? What do the pairs of stimuli that are difficult to distinguish sound like? Why might some pairs be easier to tell apart than others?

4. Compare your discrimination function with the "chance performance" level. Were any of the pairs discriminated better than would be expected by chance? Were any pairs discriminated worse than would be expected by chance?

5. Compare your discrimination function with the typical results. Are they similar?

6. Look at the results from your second experiment. Are all pairs of vowels equally easy to distinguish? Which pair or pairs of stimuli were easiest to tell apart? Which pair or pairs of stimuli were most difficult to tell apart?

7. Compare your discrimination function with the "chance performance" level. Were any of the pairs discriminated better than would be expected by chance? Were any pairs discriminated worse than would be expected by chance?

8. Compare your discrimination function with the typical results. Are they similar?

9. Compare the discrimination functions from your two experiments. Do both of the discrimination functions show that the discriminability of the sounds changes as the acoustic parameter in the stimulus continuum changes?

10. Are the shapes of the "typical" discrimination functions similar? How would you describe the difference between the two discrimination functions?

Speech Perception: Modes of Perception

1. Look at the results from your experiment. Are all pairs of notes equally easy to distinguish? Which pair or pairs of notes were easiest to tell apart? Which pair or pairs of notes were most difficult to tell apart?

2. Compare your discrimination function with the "chance performance" level. Were any of the pairs discriminated better than would be expected by chance? Were any pairs discriminated worse than would be expected by chance?

3. Compare your discrimination function with the typical results. Are they similar?

4. Look at the results from your discrimination experiments on tones and vowels. Are the shapes of the discrimination functions similar? How are they different?

5. Look at the results from your discrimination experiments on tones and consonants. Are the shapes of the discrimination functions similar? How are they different?

Speech Perception: Explaining Categorical Perception

The Motor Theory of speech perception claims that what listeners actually perceive are the intended articulatory gestures of the speaker, and thus that speech perception and speech production are intimately linked. Speech perception occurs in a specialized mode that is part of a listener's linguistic capabilities. Since there are no articulatory gestures "intermediate" between (for example) a lip-closing gesture, and a tongue tip-raising gesture, no speech sound "intermediate" between a /b/ and a /d/ can be perceived. Categorical perception arises as a consequence of the intimate connection between articulatory gestures and perception.

The Learning Theory claims that extensive experience with our own language leads us to learn which acoustic differences are relevant to the meaning of words. We thus learn to pay attention only to relevant differences and to ignore irrelevant ones (e.g., those due to differences between speakers or differences in speech rate). Once we have learned to ignore them, the irrelevant differences become hard to hear because we are not conscious of ignored differences. Since changing the duration of aspiration after the [p^h] does not affect the meaning of [p^hɛn], we ignore the duration of aspiration beyond a certain point. Categorical perception arises as a consequence of this selective attention.

The Auditory Sensitivity Theory claims that phonetic classification depends directly on the properties of the auditory system. Phonetic boundaries in stimulus continua correspond to naturally-occurring discontinuities in the (auditory) perception of similar acoustic continua. For instance, if at least 16 milliseconds must separate the onset of a noise from the onset of a buzz for listeners to be able to perceive the noise, then we expect that at least 16 milliseconds must separate the onset of voicing from a consonant release burst for listeners to be able to perceive the voicing. Categorical perception arises as a consequence of the auditory sensitivity to this and other acoustic differences.

1. What is the Motor Theory explanation for the fact that vowels appear to be perceived continuously?

2. What is the Motor Theory explanation for the fact that consonants appear to be perceived categorically?

3. What is the Learning Theory explanation for the fact that consonants appear to be perceived categorically?

4. What is the Auditory Sensitivity Theory explanation for the fact that vowels appear to be perceived continuously?

5. What is the Auditory Sensitivity Theory explanation for the fact that consonants appear to be perceived categorically?

Speech Perception: Testing the Theories

1. In the speech continuum, the onset of voicing ranges from 45 ms before the burst in the first stimulus (leftmost on the screen) to 45 ms after the burst in the last (rightmost). The intermediate stimuli change the voice onset time relative to the consonant burst in 15 ms steps. What does the first stimulus sound like? What does the last sound like? What do the intermediate stimuli sound like? Can you distinguish any of the adjacent pairs of stimuli?

2. In the non-speech continuum, the onset of one tone (at 1500 Hz) ranges from 45 ms before the onset of the second (at 500 Hz) to 45 ms after. The intermediate stimuli change the onset time of the 1500 Hz tone relative to the onset of the 500 Hz tone in 15 ms steps. Can you discriminate between the adjacent pairs of stimuli? Are all pairs in the continuum equally discriminable?

3. Are the pairs you can discriminate in the non-speech continuum analogous to the pairs you can discriminate in the speech continuum? How does your ability to discriminate between the non-speech stimuli differ from your ability to discriminate between the speech stimuli? Is the ability to discriminate between stimuli in one continuum related to the ability to discriminate between the stimuli in the other continuum?

4. When sounds are perceived continuously, in general all the pairs of stimuli are equally discriminable, whereas when they are perceived categorically, some pairs are much easier to discriminate between than others. Did you think that the stimuli in the non-speech continuum were discriminated in the way that continuously or categorically perceived stimuli are?

5. Is the way in which the non-speech stimuli are perceived consistent with Motor Theory? Is there any articulatory gesture to be perceived in the non-speech continuum?

6. Is the way in which the non-speech stimuli are perceived consistent with Learning Theory? Do you think that listeners could learn to perceive these stimuli categorically?

7. Is the way in which the non-speech stimuli are perceived consistent with Auditory Sensitivity Theory?

8. The chinchillas were trained to run to one end of the cage when they heard the first stimulus in the continuum ("ba" to English-speakers) and to the other end of the cage when they heard the last stimulus ("pa"). Were their responses to the other stimuli categorical?

9. Is the way in which the chinchillas responded to the stimuli consistent with Motor Theory? Do the chinchillas know anything about speech articulation?

10. Is the way in which the chinchillas responded to the stimuli consistent with Learning Theory, since they were not trained to respond to the intermediate stimuli?

11. Is the way in which the chinchillas responded consistent with Auditory Sensitivity Theory? Would knowing that chinchillas are mammals with hearing abilities roughly similar to humans' hearing abilities help you to decide?

12. Musicians trained in Western-style music learn to differentiate groups of simultaneous notes ("chords") and musical scales into "major" and "minor". Did the musically experienced listeners perceive these chords categorically? The inexperienced listeners?

13. Is the way in which the experienced and inexperienced listeners responded to the stimuli consistent with Motor Theory? Is any articulation involved?

14. Is the way in which the experienced and inexperienced listeners responded to the stimuli consistent with Learning Theory?

15. Is the way in which the experienced and inexperienced listeners responded to the stimuli consistent with Auditory Sensitivity Theory?

16. Is any one of these theories consistent with your interpretation of the results of these experiments? Which of the theories is consistent with most of the results? Which of the theories is inconsistent with most of the results?

17. What results would you have had to get in these experiments in order for all of them to be compatible with the Motor Theory?

18. What results would you have had to get in these experiments in order for all of them to be compatible with the Learning Theory?

19. What results would you have had to get in these experiments in order for all of them to be compatible with the Auditory Sensitivity Theory?

Speech Perception: Child and Adult Acquisition

1. Japanese does not have the /ɹ/ and /l/ phonemes found in American English; instead, it has a single phoneme that can be transcribed as /ɖ/ or /ɻ/ (a postalveolar stop or lateral approximant; Okada 1991). Therefore, Japanese speakers are not exposed to /ɹ/ and /l/ in speech. What pattern of results would you expect Japanese-speaking listeners to have in an experiment in which they tried to identify stimuli belonging to a synthetic /ɹɑ/ - /lɑ/ continuum?

2. What discrimination function would you expect Japanese-speaking listeners to have in an experiment in which they tried to discriminate between pairs of stimuli belonging to a synthetic /ɹɑ/ - /lɑ/ continuum?

3. How would you explain the difference in the discrimination functions obtained from English-speaking and Japanese-speaking subjects in the experiments that tried to discriminate between pairs of stimuli belonging to a synthetic /ɹɑ/ - /lɑ/ continuum?

4. American English does not have the /ɖ/ and /ḍ/ phonemes found in Hindi; instead, it has a single phoneme that can be transcribed as /d/. Therefore, American English speakers are not exposed to /ɖ/ and /ḍ/ in speech. What results would you expect English-speaking listeners to have in an experiment in which they tried to identify stimuli belonging to a synthetic /ɖɑ/ - /ḍɑ/ continuum?

5. What pattern of results would you expect English-speaking listeners to have in an experiment in which they tried to discriminate between pairs of stimuli belonging to a synthetic /ɖɑ/ - /ḍɑ/ continuum?

6. Both languages have /b/ phonemes. How would you explain the similarity in the ability of English-speaking and Hindi-speaking subjects to discriminate between a synthetic /bɑ/ and a synthetic /dɑ/?

7. How would you explain the difference in the ability of English-speaking and Hindi-speaking subjects to discriminate between a synthetic /ɖɑ/ and a synthetic /dɑ/?

8. How would you explain the fact that both the English-speaking and Hindi-speaking subjects discriminate between two synthetic /ɖɑ/s only slightly better than 50% of the time (i.e., only slightly better than chance)?

9. At birth, infants have not heard any of the sounds /b/, /ɖ/, or /d/. If infants' ability to discriminate between speech sounds is analogous to adults' ability to discriminate between speech sounds, would you expect young infants to be able to discriminate between synthetic /b/ and /d/? /d/ and /ɖ/? two (slightly different) synthetic /ɖ/s? (Base your answer on the pattern of results obtained from adults.)

10. Do the experimental results show that 6-8 month-old infants discriminate between synthetic /b/ and /d/ better than chance? Between /d/ and /ɖ/? Between two synthetic /ɖ/s?

11. Do the experimental results show that 11-13 month-old children growing up in English-speaking homes discriminate between synthetic /b/ and /d/ better than chance? Between /d/ and /ɖ/? Between two synthetic /ɖ/s?

12. Are the results of the experiments with 6-8 month-old infants and 10-13 month-old children as you would have expected, based on the English-speaking adults' ability to discriminate between these speech sounds? How do they differ from what you would have expected?

13. Mandarin Chinese does not have word-final /t/ and /d/ phonemes (although it does have word-initial /t/ and /t^h/, but not /d/). Therefore, Mandarin Chinese speakers are, by and large, not exposed to word-final /t/ and /d/ in speech. What pattern of results would you expect Mandarin Chinese-speaking listeners to have in an experiment in which they tried to identify word-final /t/s and /d/s?

14. Do the experimental results show that Mandarin Chinese-speaking listeners are able to identify word-final /t/ and /d/ when these speech sounds are pronounced as released [t] and [d]?

15. Do the experimental results show that Mandarin Chinese-speaking listeners are able to identify word-final /t/ and /d/ when these speech sounds are pronounced as unreleased [t̚] and [d̚]?

16. Are these results what you would have expected, based on experiments with English-speaking listeners discriminating between /d/ and /ɖ/? How do they differ from what you would have expected?

17. Does feedback (informing the listeners whether or not they just heard two different speech sounds) improve the ability of English-speaking listeners to discriminate between /d/ and /ɖ/? If so, how much does their ability improve?

18. Does feedback (informing the listeners which of the two speech sounds they just heard) improve the ability of Mandarin Chinese speaking-listeners to identify word-final /t/s and /d/s? If it does, by how much does their ability improve?

19. Does training subjects by giving them feedback during the experiment improve their ability to identify or discriminate between speech sounds that do not occur in their native language? Do the subjects' abilities improve to be as good as the identification and discrimination abilities of native speakers of the languages that do have these speech sounds?

Vowel Perception: Phonetic and Perceptual Information

1. Can you tell what vowel you are hearing when you hear just F_1? Does F_1 alone provide enough information for you to identify a vowel? To distinguish it from other vowels?

2. Do vowels with just F_1 sound natural?

3. Can you tell what vowel you are hearing when you hear F_1 and F_2? Do F_1 and F_2 together provide enough information for you to identify a vowel?

4. Do vowels with just F_1 and F_2 sound natural?

5. Can you tell what vowel you are hearing when F_3 is added? F_4? Do these formants add enough information for you to identify the vowels?

6. Do vowels with more formants sound natural? How many formants are required for a vowel to sound natural? If you were designing a speech synthesizer for a handicapped person, would you emphasize the intelligibility or the naturalness of the speech produced by the synthesizer? How many formants would your speech synthesizer have to have in order to fulfill this goal?

7. With just F_1, can you tell whether a vowel is high or low? Can you tell whether it is front or back? Rounded or unrounded? Does F_1 alone provide information about these phonetic properties of vowels?

9. With just F_2, can you tell whether a vowel is high, low, front, back, rounded or unrounded? Does F_2 alone provide information about these phonetic properties of the vowels?

10. Try various combinations of F_1 and F_2, F_3, or F_4. Can you identify vowels? Can you distinguish vowels from each other? Which formant or formants must be present in order for you to be able to distinguish the vowels in each pair from each other?

11. Suppose you were designing a cochlear implant for profoundly hearing-impaired people. Miniaturization is important, so you want the implant to transmit acoustic information to the cochlea as efficiently as possible. What is the minimum set of formants that must be transmitted in order for the implant user to be able to identify vowels unambiguously?

12. Some cochlear implants have extracted the frequencies of the second and third formants and transmitted those to the cochlea. Do you think this was a good design decision? Why?

Vowel Perception: Normalization

1. Some researchers have suggested that scaling the formant frequencies of a speaker's vowels by a constant ratio can normalize them. What is the ratio between the child's F_1 in /i/ and the adult's? What is the ratio between the child's F_2 in /i/ and the adult's? Are these ratios the same?

2. What is the ratio between the child's F_1 in /æ/ and the adult's? What is the ratio between the child's F_2 and the adult's in /æ/? What are the ratios in /ɑ/? /u/?

3. Are these ratios the same? Close to the same? Is the suggestion that there is a single ratio that characterizes the differences between the child's vowels and the adult's plausible? Why?

3. The suggestion that a single ratio can be used to normalize the formant frequencies of a child's vowels seems reasonable since the formant frequencies of many vowels correspond to wavelength resonances of the vocal tract which depend on the length of the vocal tract. Which vowels have formant frequencies that correspond to wavelength resonances of the vocal tract?

4. However, some vowels have formants that correspond to cavity (Helmholtz) resonances. What does the resonant frequency of a cavity depend on? Would you expect the frequency of a formant that corresponds to a cavity resonance of the vocal tract to depend directly on the length of the vocal tract?

5. Which vowels do you think have cavity resonances?

6. Would you expect the formant frequencies of the vowels with wavelength resonances to have the same normalizing ratio as the vowels with cavity resonances? Why?

7. Some researchers have suggested that each vowel, no matter who says it, has a characteristic specific ratio of formant frequencies. What is the ratio between F_1 and F_2 in the child's /i/? In the adult's /i/? Are these ratios the same?

8. What is the ratio between F_1 and F_2 in the child's /æ/? /ɑ/? /u/? The adult's?

9. Does each vowel have a particular ratio between F_1 and F_2? Is the suggestion that each vowel is characterized by a specific ratio of formant frequencies plausible? Why?

10. In vowels with formants that correspond to wavelength resonances, the formant frequencies are proportional to the lengths of the front and back cavities of the vocal tract (formed by moving the tongue around). How do different-sized vocal tracts have to be related to each other in order to keep the ratio between F_1 and F_2 constant?

Vowel Perception: Auditory Processing

1. Examine the acoustic spectrograms and the auditory spectrograms of the various vowels. Is F_1 shown in both spectrograms? Is F_2 shown in both? Is F_3 shown in both? Which formants are distinct in the acoustic spectrogram? Which formants are distinct in the auditory spectrogram?

2. Compare the auditory spectrograms of the child's and adult's vowels. Is F_1 clearly represented in the spectrograms? Is F_2 clearly represented? F_3?

3. Are the child's vowels and the adult's vowels represented similarly in the auditory spectrograms? What feature of the auditory spectrogram affects how the vowels are shown?

Vowel Perception: Auditory Experiments

1. What are the frequencies of F_2 and F_3 in the model /ɛ/?

2. What is the frequency of F_2 that you picked in the matching vowel? Is it the same as the frequency of F_2 in the model /ɛ/? As F_3? What is the relation between the frequency of F_2 that you picked and the formant frequencies of the model /ɛ/?

3. Does the F_2 that you picked seem closer to the F_2 in the model /ɛ/ in the acoustic spectrogram or in the auditory spectrogram? Which kind of spectrogram better represents the perceptual similarity of the model /ɛ/ to the matching vowel?

1. What are the frequencies of F_1 and F_2 in the model /ɔ/?

2. What is the frequency of F_1 that you picked in the matching vowel? Is it the same as the frequency of F_1 in the model /ɔ/? As F_2? What is the relation between the frequency of F_1 that you picked and the formant frequencies of the model /ɔ/?

3. Are the formant frequencies you picked to match the model /ɛ/ and to match the model /ɔ/ related to the formant frequencies of the models in the same way? What conclusions do you draw?

Acknowledgments

The IPA Kiel font used in this work is available from Linguist's Software, Inc., P.O. Box 580, Edmonds WA 98020-0580, tel (206)775-1130. Multimedia ToolBook and ToolBook are trademarks of Asymetrix Corporation. Microsoft, Windows, Windows 95, and Windows NT are trademarks of Microsoft Corporation. All other brand names or product names may be trademarks or registered trademarks of their respective owners.